CONDITIONAL PROPERTIES OF JAMES-STEIN SET ESTIMATORS

GEORGE CASELLA[1]

Cornell University

BU-856-M September, 1984

[1] Research supported by National Science Foundation Grant No.MCS83-00875.

AMS 1980 *subject classifications.* Primary 62C99; secondary 62F10.

Key words and Phrases. Betting procedures, confidence sets, multivariate normal.

Summary

The usual confidence sphere for a multivariate normal mean can be uniformly improved upon, it terms of coverage probability, by recentering it at a James-Stein positive part estimator. Furthermore, there is strong evidence that uniform improvement in terms of both volume and coverage probability can also be attained. However, these improved sets have been criticized because of suspected poor conditional performance. Using the theory of relevant betting procedures, which provides an objective means for assessing the conditional performance of a statistical procedure, it is shown that a large class of James-Stein confidence sets exhibit good conditional behavior: they do not allow the existence of super-relevant betting procedures.

CONDITIONAL PROPERTIES OF JAMES-STEIN SET ESTIMATORS

By GEORGE CASELLA[1]

Cornell University

1. **Introduction.** Recent advances in the theory of set estimation have led to some disturbing questions within the frequentist theory of statistics. In particular, if C(X) is a set estimator for a parameter θ that satisfies

$$P_\theta[\theta \in C(X)] > 1-\alpha \qquad \text{for all } \theta, \tag{1.1}$$

then one might question the precision of asserting that one is "$1-\alpha$ confident" in the set estimator C(X). Instead, we would rather assert some confidence level greater that $1-\alpha$ since we know that, given (1.1), the coverage probability at any value of θ will be greater than $1-\alpha$. Confidence sets with property (1.1) have received renewed criticism of late, one of most recent from Hinkley (1983):

> "...I do not find it useful to attach the probability .9, say, to an interval that covers with probability 1. The interval is only useful if we can state what the probability is, with reasonable accuracy."

Concerns such as these expressed by Hinkley are not new; Stevens (1950), in a paper dealing with binomial confidence intervals states, "<u>...it is a statistician's duty to be wrong</u> the stated proportion of times." The criticisms of both Hinkley and Stevens can, of course, be applied to any

---

[1]Research supported by National Science Foundation Grant No.MCS83-00875.

AMS 1980 *subject classifications.* Primary 62C99; secondary 62F10.

Key words and phrases. Betting procedures, confidence sets, multivariate normal.

conservative procedure, and point out that the user of a procedure that satisfies (1.1) might be "cheated" by only being allowed to assert a confidence level of $1-\alpha$.

Part of the problem is that frequentist measures of precision are pre-experimental measures. Frequentist inference is unconditional (with respect to the realized data), so it cannot have the data-dependent precision that either Bayesian or likelihood inference has. Thus there is a trade off between a more precise, but conditional, inference and a less precise, but unconditional, inference.

Another problem is that classical frequentist theory seems to try to be too objective, guarding against values of θ that are unreasonable (e.g., $\theta = \infty$). It is possible to make frequentist statements more precise by limiting the range of θ in some way, which is a basic idea underlying the "parametric empirical Bayes" approach of Morris (1983). Another "way out" of this problem is to try to combine the more precise Bayesian conclusions with some degree of frequentist objectivity, which is an underlying idea of the "robust Bayesian" approach of Berger (1982). Both of these approaches have merit, and succeed in producing more precise inferences (at the cost of some objectivity), but they do not address the main issue with which we are concerned: Under what conditions can the frequentist feel comfortable in assigning a confidence of $1-\alpha$ to a set estimator that satisfies (1.1)?

The answer to this question, we feel, lies within the theory of relevant betting procedures, a theory which provides an objective measure of the conditional performance of a set estimator. The beginnings of the theory date back to Fisher (1956), but were first formalized by Buehler (1959). Buehler argued that one can make the statement "the probability that the set $C(X)$ covers θ is $1-\alpha$" only if one is willing to accept bets that $\theta \in C(X)$ at odds $1-\alpha:\alpha$, and accept bets that $\theta \in C(X)$ at odds $\alpha:1-\alpha$.

If, for a given set estimator $C(X)$, there exists a subset, S, of the sample space (a recognizable subset in the terminology of Fisher, 1956), that satisfies either

$$P_\theta(\theta \in C(X) | X \in S) > 1-\alpha+\varepsilon \qquad \text{for all } \theta,$$

(1.2) or

$$P_\theta(\theta \in C(X) | X \in S) < 1-\alpha-\varepsilon \qquad \text{for all } \theta,$$

for some $\varepsilon > 0$, the one should have doubts about assigning confidence $1-\alpha$ to the set $C(X)$. A subset S that satisfies (1.2) is called a relevant subset for $C(X)$, and provides a winning betting strategy against $C(X)$. Conversely, if no such subsets exist for a set estimator $C(X)$, with confidence $1-\alpha$, then there is no winning strategy and we are justified in assigning confidence $1-\alpha$ to $C(X)$ no matter what the coverage probability may be.

We will refer to the pair $\langle C(X), 1-\alpha \rangle$ as a confidence set for the parameter θ with confidence level $1-\alpha$. (In general, α may be a function of X, $\alpha=\alpha(X)$, but here we will restrict attention to a constant confidence level.) The confidence set $\langle C(X), 1-\alpha \rangle$ need not be a frequentist $1-\alpha$ confidence region in that

$$\inf_\theta P[\theta \in C(X)] = 1 - \alpha \quad , \tag{1.3}$$

but we assume that if one is willing to use $\langle C(X), 1-\alpha \rangle$, then one is willing to accept all bets in the manner stated above. If, however, $1-\alpha$ is also the frequentist confidence level in the sense of (1.3), then $\langle C(X), 1-\alpha \rangle$ has, to a certain extent, answered the criticisms of Hinkley and Stevens. The theory of relevant betting procedures, therefore, gives us a mathematical framework for evaluating the conditional performance of a frequentist confidence set.

If α satisfies (1.3) and there are no relevant subsets for $\langle C(X),1-\alpha\rangle$, then $\langle C(X),1-\alpha\rangle$ is an extremely desirable confidence set. Not only does it have the frequentist property of guaranteeing a minimum coverage probability, but it also has no serious conditional flaws. Such properties are enjoyed by the usual confidence sphere for a multivariate normal mean. The main purpose of this paper is to show that similar properties are enjoyed by confidence spheres recentered at some James-Stein type estimators (which formed part of the target of Hinkley's criticism). Moreover, these recentered spheres may even have reduced radius and still remain $1-\alpha$ confidence sets within the theory of relevant betting procedures.

2. **Betting Procedures.** Buehler's concept of relevant subsets was extended and formalized by Robinson (1979a) to the concept of relevant betting procedures (i.e., functions). Betting strategies exist which cannot be expressed in terms of subsets, so Robinson's extension was intended to include all possible betting strategies. Thus, a betting procedure, $s(X)$, is defined to be any bounded function of X. Without loss of generality we take this bound to be unity. We can think of $|s(x)|$ as the probability that a bet of one unit is made when $X=x$ is observed, with the sign of $s(X)$ giving the direction of the bet.

If $\langle C(X),\beta(X)\rangle$ is a confidence set, then the betting procedure $s(X)$ is said to be

i. semirelevant if

$$E_\theta\{[I(\theta \in C(X)) - \beta(X)] \geq 0 \qquad \text{for all } \theta$$

and is strictly positive for some θ,

ii. relevant if, for some $\varepsilon > 0$,

$$E_\theta\{[I(\theta \in C(X)) - \beta(X)]s(X)\} \geq \varepsilon E_\theta|s(X)| \qquad \text{for all } \theta$$

with strict inequality for some θ,

iii. super-relevant if, for some $\varepsilon > 0$

$$E_\theta\{[I(\theta \in C(x)\ ) - \beta(X)]s(X)\} \geq \varepsilon \qquad \text{for all } \theta.$$

Notice that if $s(X)$ is the indicator function of some set, then the above definition of a relevant betting procedure reduces to that of Buehler.

If X is an observation from a p-variate normal distribution with mean vector θ and covariance matrix I, the usual frequentist $1-\alpha$ confidence set is

$$C_0(X) = \{\theta: |\theta - X| \leq c\} \tag{2.1}$$

where c satisfies $P(\chi_p^2 < c^2) = 1-\alpha$. From the results of Robinson (1979b), it follows that $C_0(X)$ does not allow relevant betting procedures, which is the strongest conditional property that one can expect $C_0(X)$ to have. (Only proper Bayes posterior confidence sets do not allow the existence of semirelevant betting procedures.)

When evaluating procedures on both frequentist and conditional grounds, there seems to be a trade-off in the sense that procedures with the strongest conditional properties (i.e., proper Bayes procedures) have weaker frequentist properties, and vice versa. Since we are addressing conditional properties from a frequentist standpoint, and want to use procedures that are $1 - \alpha$ procedures in a frequentist sense, we cannot hope that these procedures have the strongest conditional properties. What we can hope for, instead, is that our frequentist procedures have a conditional property that is strong enough to eliminate any aberrant behavior. According to Bondar (1977), his conditional confidence procedure, which is equivalent to the absence of super-relevant betting procedures, is strong enough to do this. In particular, it is strong enough to eliminate Cox's paradox (Cox, 1958). Thus, we focus on the existence of super-relevant betting procedures.

An improvement on $C_0(X)$, in frequentist terms, is obtained by recentering $C_0(X)$ at a positive-part James Stein estimator. The set estimator

(2.2)
$$C_\delta(X) = \{\theta : |\theta - \delta(X)| \le c\} ,$$
$$\delta(X) = \left(1 - \frac{a}{|X|^2}\right)^+ X$$

satisfies

(2.3) $$P_\theta[\theta \varepsilon C_\delta(X)] > P_\theta[\theta \varepsilon C_0(X)] \quad \text{for all } \theta,$$

for a range of values of a, $0 \le a \le a_0 < p - 2$. (Hwang and Casella, 1982). Moreover, there is strong evidence, (Casella and Hwang, 1983), mostly numerical, that the set estimator

$$C_\delta^v(X) = \{\theta : |\theta - \delta(X)| \le cv(X)\} \tag{2.4}$$

also dominates $C_0(X)$ in coverage probability for certain choices of $v(X)$, $|v(X)| \le 1$.

The conditional criticisms voiced against $C_\delta(X)$ are cause for concern, but only minor concern, since $C_\delta(X)$ will tend to err on the conservative side; covering θ more often than the nominal confidence level asserts. In betting terminology, $\langle C_\delta(X), 1-\alpha\rangle$ allows the existence of a <u>positively biased</u> semirelevant betting procedure, since $s(X) = 1$ will give $E_\theta\{[I(\theta \varepsilon C_\delta(X)] - (1-\alpha)\}s(X) \ge 0$. In practical applications this deficiency is extremely minor: we tend to forgive errors on the conservative side. Criticisms of $C_\delta^v(X)$ are more serious, since it seems possible to make $v(X)$ very small over a range of X values, leading to confidence sets that, conditionally, will cover θ with probability much less than $1-\alpha$. This criticism is more serious because it cast doubts on whether the frequentist $1-\alpha$ guarantee is really guaranteeing anything.

These conditional criticisms are answered, within the theory of relevant betting procedures, in the next section. The answer is quite simple: under mild conditions on $\delta(X)$ and $v(X)$, neither $\langle C_\delta(X), 1-\alpha\rangle$ nor $C_\delta^v(X)$, $1-\alpha\rangle$ allow the existence of super-relevant betting procedures.

3. **James-Stein Confidence Sets.** The set estimator $C_\delta(X)$ of (2.2) is a frequentist $1-\alpha$ confidence set. The next theorem shows that it is also reasonable to regard $\langle C_\delta(X), 1-\alpha\rangle$ as a $1-\alpha$ confidence set within the theory of relevant betting procedures.

Theorem 3.1: Let $X \sim N_p(\theta, I)$, $p \geq 3$, and let α and c satisfy $P_\theta(|X - \theta| \leq c) = P(\chi_p^2 \leq c^2) = 1-\alpha$. Let $\delta(X) = [1 - \gamma(|X|)]X$, $0 \leq \gamma(X) \leq 1$, and let

$$C_\delta(X) = \{\theta : |\theta - \delta(X)| \leq c\} \quad .$$

If there exist positive constants K_0, K_1 and $r > 1$ such that $|\gamma(|X|)| < K_1/|X|^r$ for $|X|^2 > K_0$, then there are no super-relevant betting procedures against the set estimator $\langle C_\delta(X), 1-\alpha\rangle$.

Proof: Suppose $s(X)$ is super-relevant for $\langle C_\delta(X), 1-\alpha\rangle$. Then there exist $\varepsilon > 0$ such that

$$E_\theta\{I[\theta \varepsilon C_\delta(X)] - (1-\alpha)\}s(X) \geq \varepsilon \tag{3.1}$$

for all θ. Multiply both sides of (3.1) by $\pi_b(\theta)$, a $N[0, (b^{-1} - 1)]$ density $(0 < b < 1)$ and integrate over all θ. If $s(X)$ is a super-relevant betting procedure it then follows that

$$\int_\Theta E_\theta\{I[\theta \varepsilon C_\delta(X)] - (1 - \alpha)\}s(X)\pi_b(\theta)d\theta \geq \varepsilon \ . \tag{3.2}$$

The proof will proceed by showing that for sufficiently small b the left hand side of (3.2) is less than ε. This will contradict the supposition that $s(X)$ is super-relevant and establish the theorem. To this end, write

$$\begin{aligned} &\int_\Theta E_\theta\{I[\theta \varepsilon C_\delta(X)] - (1 - \alpha)\}s(X)\pi_b(\theta)d\theta \\ &\quad = \int_X \Bigg[\int_{\theta \varepsilon C_\delta(X)} \pi_b(\theta|X)d\theta - (1-\alpha)\Bigg]s(X)m_b(X)dX \ , \end{aligned} \tag{3.3}$$

when $\pi_b(\theta|X)$ is the conditional density of θ given X, $N[(1-b)X,(1-b)I]$, and $m_b(X)$ is the marginal density of X, $N(0,b^{-1}I)$.

The integration over X in (3.3) will be split into three pieces: let h be a positive constant, satisfying $\frac{1+h}{1-h} < r$, and consider the three sets $\{|X|:|X|^2 < b^{-1+h}\}$, $\{|X|:|X|^2 > b^{-1-h}\}$, and $\{X:b^{-1+h} < |X|^2 < b^{-1-h}\}$. Note that the union of these three sets is the entire region of integration.

Consider first the integration over the set $\{S:|X|^2 < b^{-1+h}\}$. We have

$$(3.4)\qquad \int_{\{X:|X|^2< b^{-1+h}\}} \left[\int_{\theta\varepsilon C_\delta} \pi_b(\theta|X)d\theta - (1-\alpha)\right]s(X)m_b(X)dX$$

$$(3.5)\qquad \begin{aligned} &\le \int_{\{X:|X|^2< b^{-1+h}\}} |s(X)|m_b(X)dX \\ &\le P_b(|X|^2 < b^{-1+h}) \end{aligned}$$

since $|s(X)| \le 1$. Furthermore,

$$(3.6)\qquad P_b(|X|^2 < b^{-1+h}) = P_b(b|X|^2 < b^h) = P(\chi_p^2 < b^h) \to 0 \quad \text{as } b \to 0$$

The integration over $\{X:|X|^2 > b^{-1-h}\}$ is handled in the same manner. The contribution to (3.3) from this region is bounded by $P_b(|X|^2 > b^{-1-h})$, which also goes to zero as $b\to 0$.

It only remains to establish that the integral over the middle region can be made sufficiently small. If we show that, for sufficiently small b,

$$(3.7)\qquad \int_{\{X:b^{-1+h}< |X|^2 < b^{-1-h}\}} \left[\int_{\theta\varepsilon C_\delta(X)} \pi(\theta|X)d\theta - (1-\alpha)\right]s(X)m_b(X)dX \le (\varepsilon/3)$$

it will then follow that (3.2), and hence (3.1) is contradicted, and the theorem will be established.

Straight forward transformations will establish that

$$\int_{\theta\varepsilon C_\delta(X)} \pi_b(\theta|X)d\theta = \frac{1}{[2\pi(1-b)]^{\frac{1}{2}}} \int_{|\delta(X)|-c}^{|\delta(X)|+c} P\left(\chi^2_{p-1} \le \frac{1}{1-b}\{c^2 - [y-|\delta(X)|]^2\}\right) \tag{3.8}$$
$$\times\ e^{-\frac{1}{2}\frac{[y-(1-b)|X|]^2}{(1-b)}} dy \quad .$$

Using the facts that $|\delta(X)|-c \le y \le |\delta(X)| + c, 0 \le \gamma(X) \le 1, |\gamma(X)| < K_1/|X|^r$ for $|X|^2 > K_0$, $b^{-1+h} < |X|^2 < b^{-1-h}$ and $r > (1+h)/(1-h)$, it can be shown that for suffiently small b,

$$(y - |\delta(X)|)^2 - 0(b) \le \frac{[y - (1-b)|X|]^2}{1-b} \le [y - |\delta(X)|]^2 + 0(b), \tag{3.9}$$

where $0(b)\to 0$ as $b\to 0$. Also, for $|\delta(X)| - c < y < |\delta(X)| + c$,

$$P\left(\chi^2_{p-1} \le \{c^2 - [y-|\delta(X)|]^2\}\right) \le P\left(\chi^2_{p-1} \le \frac{1}{1-b}\{c^2- [y-|\delta(X)|]^2\}\right)$$
$$\le P\left(\chi^2_{p-1} \le \{c^2-[y-|\delta(X)|]^2\}\right) \tag{3.10}$$
$$+ P\left\{c^2-[y-|\delta(X)|]^2 \le \chi^2_{p-1} \le c^2-[y-|\delta(X)|^2] + \frac{bc^2}{1-b}\right\}$$

Using the fact that

$$\max_{a:a>0} P(a \le \chi^2_{p-1} \le a + \varepsilon) = \begin{cases} P(\chi^2_{p-1} \le \varepsilon) & \text{if } p = 3 \\ P[a^*(\varepsilon) \le \chi^2_{p-1} \le a^*(\varepsilon)+\varepsilon] & \text{if } p > 3, \end{cases}$$

where $a^*(\varepsilon) = \varepsilon\exp[- \varepsilon/(p - 3)]/\{1 - \exp[-\varepsilon/(p - 3)]\}$, it follows that, for $|\delta(X)| - c < y < |\delta(X)| + c$.

$$P\left\{c^2 - [y - |\delta(X)|]^2 \le \chi^2_{p-1} \le c^2 - [y - |\delta(X)|^2] + \frac{bc^2}{1-b}\right\} = 0(b). \tag{3.11}$$

Finally, noting that

$$\frac{1}{\sqrt{2\pi}} \int_{|\delta(X)|-c}^{|\delta(X)|+c} P\{\chi^2_{p-1} \le c^2- [y - |\delta(X)|^2]\}e^{-\frac{1}{2}[y-|\delta(X)|]^2}dy = 1 - \alpha , \tag{3.12}$$

we have, using (3.8), (3.9), (3.10), (3.11) and (3.12),

$$\frac{e^{-O(b)}(1-\alpha)}{(1-b)^{\frac{1}{2}}} \le \int_{\theta\varepsilon C_\delta(X)} \pi_b(\theta|X)d\theta \le \frac{e^{O(b)}(1-\alpha)}{(1-b)^{\frac{1}{2}}} + O(b)$$

for $b^{-1+h} < |X|^2 < b^{-1-h}$. So, for sufficiently small b,

$$\int_{\{X:b^{-1+h}<|X|^2<b^{-1-h}\}} \left[\int_{\theta\varepsilon C_\delta(x)} \pi_b(\theta|X)d\theta - (1-\alpha)\right] s(X)m_b(X)dX$$

$$\le \int_{\{X:b^{-1+h}<|X|^2<b^{-1-h}\}} \left|\int_{\theta\varepsilon C_\delta(X)} \pi_b(\theta|X)d\theta - (1-\alpha)\right| |s(X)|m_b(X)dX$$

$$< (\varepsilon/3) \int_{\{X:b^{-1+h}<|X|^2<b^{-1-h}\}} |s(X)|m_b(X)dX \tag{3.15}$$

$$\le (\varepsilon/3)P(b^{-1+h} < |X|^2 < b^{-1-h})$$

$$\le \varepsilon/3$$

completing the proof. ||

In retrospect, the result in Theorem 3.1 is not all that surprising; the fact that there are no super-relevant betting procedures against $\langle C_\delta(X), 1-\alpha\rangle$ is closely related to the fact that none exist against $\langle C_0, 1-\alpha\rangle$. This is somewhat analogous to the case of minimax point estimation, where all minimax estimators must collapse to X as $|X|\to\infty$.

Theorem 3.1 can be extended to cover the more general confidence set $\langle C^v_\delta(X), 1-\alpha\rangle$.

Theorem 3.2: Suppose $\delta(X)$ and $\gamma(|X|)$ are as in Theorem 3.1, and define

$$C^v_\delta(X) = \{\theta: |\theta - \delta(X)| \le cv(X)\} ,$$

where c satisfies $P(\chi^2_p \le c^2) = 1-\alpha$ and $v(X) > 0$ for all X. If there exist constants K_1, K_2 and $t>0$ such that $|v(X)-1| \le K_2/|X|^t$ for $|X|^2 > K_1$, then there are no super-relevant betting procedures against the set estimator $\langle C_\delta^v(X), 1-\alpha\rangle$.

<u>Proof</u>: The proof of Theorem 3.1, with minor modifications, will also suffice to prove this theorem. We will only indicate the modifications.

Clearly, the arguments used for the regions $\{X: |X|^2 < b^{-1+h}\}$ and $\{X: |X|^2 > b^{-1-h}\}$ will also serve here, so we only need concentrate on $\{X: b^{-1+h} < |X|^2 < b^{-1-h}\}$. The counterpart to equation (3.8) is

$$\int_{\theta\varepsilon C_\delta^v(X)} \pi_b(\theta|X)d\theta = \frac{1}{[2\pi(1-b)]^{\frac{1}{2}}} \int_{|\delta(X)|-cv(X)}^{|\delta(X)+cv(X)} P\left(\chi^2_{p-1} \le \frac{1}{1-b}\{c^2v^2(X)-[y-|\delta(X)|^2]\}\right) \tag{3.15}$$
$$\times\, e^{-\frac{1}{2}\frac{[y-(1-b)|X|]^2}{(1-b)}}\, dy \quad ,$$

and the assumption on $v(X)$ is enough to show that, for $y\varepsilon|\delta(X)| \pm cv(X)$,

$$\frac{[y-|\delta(X)|]^2}{v(X)} - 0(b) \le \frac{[y-(1-b)|X|]^2}{1-b} \le \frac{[y-|\delta(X)|]^2}{v(X)} + 0(b) \quad . \tag{3.16}$$

The argument used in establishing (3.11) can also be used to establish

$$P\left(c^2v^2(X) - (y-|\delta(X)|)^2 \le \chi^2_{p-1} \le c^2v^2(X) - (y-|\delta(X)|)^2 + \frac{bc^2}{1-b}\right) = 0(b).$$

Since we also have

$$\frac{1}{\sqrt{2\pi}} \int_{|\delta(X)|-cv(X)}^{|\delta(X)|+cv(X)} P\{\chi^2_{p-1} \le c^2v^2(X)-[y-|\delta(X)]^2\} e^{-\frac{1}{2}\frac{[y-|\delta(X)|]^2}{v(X)}} dy = 1-\alpha + 0(b) \tag{3.17}$$

bounds analagons to (3.13) can be obtained, establishing the theorem. ||

<u>Remark</u>: The condition $v(X) > 0$ for all X is crucial in obtaining the bounds in (3.16). Clearly, if $v(X) = 0$ for a set of positive Lebesgue measure, then the set $\{X: v(X) = 0\}$ forms the basis for a super-relevant betting procedure against $\langle C_\delta^v(X), 1-\alpha\rangle$.

The conditions of Theorems 3.1 and 3.2 cover confidence sets centered at the positive-part James-Stein estimator, but not those centered at the ordinary James-Stein estimator. Confidence sets centered at the ordinary James-Stein estimator.

$$(3.18) \qquad C_{\delta^{JS}}(X) = \{\theta : |\theta - \delta^{JS}(X)| \le c\} \quad , \quad \delta^{JS}(X) = \left(1 - \frac{a}{|X|^2}\right)X,$$

do allow the existence of super-relevant betting procedures, as the following argument shows.

The coverage probability of $C_{\delta^{JS}}(X)$ can be calculated by integrating over the region

$$(3.19) \qquad \left\{|X|, \beta : |\theta|^2 - 2|\theta|\left(1 - \frac{a}{|X|^2}\right)|X|\cos\beta + \left(1 - \frac{a}{|X|^2}\right)^2 |X|^2 \le c^2\right\} \quad ,$$

where $\cos\beta = \theta'X/|\theta||X|$. It is easy to see that this region is contained in

$$(3.20) \qquad \left\{|X| : \left[\,|\theta| - \left(1 - \frac{a}{|X|^2}\right)|X|\,\right]^2 \le c^2\right\} \quad .$$

Consider the intersection of the region in (3.20) with $\{|X| : |X|^2 < h\}$, where $0 < h < a$. If $|\theta| > c$, this intersection is empty. If $|\theta| \le c$ the intersection is given by the region

$$(3.21) \; \{|X| : r_-(|\theta|) < |X| < h^{\frac{1}{2}}\} \; , \; r_-(|\theta|) = \tfrac{1}{2}\{|\theta| - c + [(|\theta| - c)^2 + 4a]^{\frac{1}{2}}\}.$$

Since $r_-(|\theta|) \ge r_-(0) = \frac{1}{2}[-c + (c^2 + 4a)^{\frac{1}{2}}] > 0$ for $0 \le |\theta| \le c$, the intersection is empty for all $|\theta|$ if $h^{\frac{1}{2}} < r_-(0)$. Therefore, the betting procedure $s(X) = -I_{[0, r_-(|\theta|)]}(|X|^2)$ is super-relevant for $C_{\delta^{JS}}(X)$. (This betting procedure bets against coverage if $|X|$ is small.) Note that this conclusion holds no matter what confidence level we attach to $C_{\delta^{JS}}(X)$.

The key defect in $\delta^{JS}(X)$ is the unboundedness of $|\delta^{JS}(X)|$ for finite X. Modifications of the above argument can be used to show the existence of super-relevant betting procedures for any estimator displaying this behavior.

4. **Comments.** Critics of confidence sets such as $\langle C_\delta(X), 1-\alpha\rangle$ and $\langle C_\delta^V(X), 1-\alpha\rangle$ will, no doubt, not be completely mollified by the results presented here. Once can argue, for example, that the non-existence of super-relevant betting procedures is really not a positive property, but rather the absence of a negative property; somewhat akin to the property of admissibility. (An admissible estimator is not necessarily good, but an inadmissible estimator can be improved.) However, if we regard this property as we regard admissibility, then we should choose our candidate estimators only from those that have the property. Under this principle, some James-Stein set estimators are reasonable alternatives to $C_0(X)$.

Furthermore, all the confidence set considered here, including $C_0(X)$, allow the existence of semirelevant betting procedures. Thus, one can find a betting procedure for which the expected return is non-negative. While this less serious conditional flaw still may be a basis for criticism, it can be forgiven on two counts. First, we are considering alternatives to $C_0(X)$, and it is reasonable to look among procedures whose conditional performance is similar to that of $C_0(X)$, and not necessarily stronger. Second, it seems quite unlikely that any procedure, other than a proper Bayes posterior set, will allow the existence of semirelevant betting procedures. Therefore, if there is desire to retain some objectivity in the frequentist sense, it seems that one will have to put up with semirelevant procedures.

REFERENCES

BERGER, J. O. (1982). The robust Bayesian viewpoint. Technical Report #82-9, Purdue University.

BONDAR, J. V. (1977). A conditional confidence principle. *Ann. Statist.* **5** 881-891.

BUEHLER, R. J. (1959). Some validity criteria for statistical inferences. *Ann. Math. Statist.* **30** 866-877.

CASELLA, G. and HWANG, J. T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution. *J. Amer. Statist. Assoc. – Theory and Methods Section* **78** 688-698.

COX, D. R. (1958). Some problems connected with statistical inference. *Ann. Math. Statist.* **29** 357-372.

FISHER, R. A. (1956). *Statistical methods and scientific inference*, Oliver and Boyd, Edinburgh.

HINKLEY, D. (1983). Comment on "Parametric empirical Bayes inference." *J. Amer. Statist. Assoc.* **78** 58-59.

HWANG, J. T. and CASELLA, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.* **10** 868-881.

MORRIS, C. (1983). Parametric empirical Bayes inference: Theory and applications. *J. Amer. Statist. Assoc.* **78** 47-55.

ROBINSON, G. K. (1979a). Conditional properties of statistical procedures. *Ann. Statist.* **7** 742-755.

ROBINSON, G. K. (1979b). Conditional properties of statistical procedures for location and scale parameters. *Ann. Statist.* **7** 756-771.

STEVENS, W. L. (1950). Fiducial limits for the parameter of a discontinuous distribution. *Biometrika* **37** 117-121.